0517.69179 PATENT APPLICATION

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In Re U.S. Patent Application)
)
Applicant: Thomas Michael Sailer)
)
Serial No. 10/763,507)
)
Filed: January 23, 2004)
)
For: FAULT TOLERANT COMPUTER CONTROLLED SYSTEM)
)
Art Unit: Unassigned)

I hereby certify that this paper is being deposited with the United States Postal Service as First Class Mail in an envelope addressed to: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450, on this date.

February 19, 2004
Date

Reg. No. 30,270
Attorney for Applicant

CLAIM FOR PRIORITY

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

Applicants claim foreign priority benefits under 35 U.S.C. 119 on the basis of the foreign application identified below:

European Patent Application No. 03001308.0

Filed: January 23, 2003

A certified copy of the priority document is enclosed.

Respectfully submitted,

GREER, BURNS & CRAIN, LTD.

By
Paul G. Juettner
Reg. No. 30,270

February 19, 2004

300 South Wacker Drive
Suite 2500
Chicago, IL 60606
(312) 360-0080
Customer No. 24978

Anmeldung Nr:
Application no.: 03001308.0
Demande no:

Anmeldetag:
Date of filing: 23.01.03
Date de dépôt:

Anmelder/Applicant(s)/Demandeur(s):

Supercomputing Systems AG
Hauptstrasse 127
4716 Welschenrohr
SUISSE
Supercomputing Systems AG
Technoparkstrasse 1
8005 Zürich
SUISSE

Bezeichnung der Erfindung/Title of the invention/Titre de l'invention:
(Falls die Bezeichnung der Erfindung nicht angegeben ist, siehe Beschreibung.
If no title is shown please refer to the description.
Si aucun titre n'est indiqué se referer à la description.)

Fault tolerant computer controlled system

In Anspruch genommene Priorität(en) / Priority(ies) claimed /Priorité(s) revendiquée(s)
Staat/Tag/Aktenzeichen/State/Date/File no./Pays/Date/Numéro de dépôt:

Internationale Patentklassifikation/International Patent Classification/
Classification internationale des brevets:

G06F11/00

Am Anmeldetag benannte Vertragstaaten/Contracting states designated at date of filing/Etats contractants désignées lors du dépôt:

AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LU MC NL
PT SE SI SK TR LI

Fault tolerant computer controlled system

The present invention relates to a fault tolerant computer controlled system as it can e.g. be used for controlling a vehicle or other critical device.

As computer systems gain increasing significance in many applications of human life, their reliability becomes more and more important because a failure may have dire consequences, including injury or casualties. Example of such computer systems are vehicle guidance or control systems, such as train guidance or aircraft control systems, as well as medical systems.

Typical "mean times between failure" of electronic computers are in the order of 10^4 hours, which is unacceptably high for critical applications. Hence, it has been common practice to use several computers in a parallel, redundant operation in order to increase reliability.

Conventional redundant systems generally use a plurality of computers, which act as data sources in a network. The network consists of a plurality of communication links, each of which connects one computer with a data receiver, such as an actuator for a flap in an aircraft. The computers generate data items containing commands for the flap's operation. The flap receives all data items and combines them for generating an error tolerant data item, e.g. by determining a median value.

This type of system is unable to transmit data items upon failure of a communication link. To overcome this, it has been suggested to interconnect the computers using additional communication links. In case a communication link between a given computer and a data receiver is found to fail, the data items from the given computer are re-routed to other computers and an alternative communication link. As systems of this type may contain a large number of computers and receivers and even a larger number of communication links, the required steps

for re-routing the data items upon failure of a communication link may become fairly complex. Also, analysis and testing of the system for all possible failures and re-routing configurations becomes very complicated and expensive if not impossible.

The problem to be solved by the present invention is therefore to provide a method and system of the type mentioned above that is easier to implement, to analyze and to test while maintaining a low risk of failure.

This problem is solved by the method and system according to the independent claims.

Hence, according to the invention, data is fed into the receiver communication links from the outputs of a switching assembly. The switching assembly has several inputs, each of which is connected to a data source or to another output. The switching assembly is able to connect each data source to each data receiver over at least two different receiver communication links. The whole system is adapted to send every data item from any given data source to any given data receiver through every one of the at least two different receiver communication links such that the given data receiver receives the same data item through at least two receiver communication links.

In normal operation, each data receiver receives every data item at least twice through separate communication links. Even though this requires additional bandwidth, it has the advantage that no re-routing of data items is required if a fault in a communication link occurs, i.e. the flow of information does not have to be rearranged when a failure occurs, which makes the system more reliable and easier to analyze and to test. It is easy to predict what kind of failures the system is able to handle and there is no need to test all possible combinations of potential failures.

In a preferred embodiment, the switching assembly is divided into a plurality of switching units, wherein each switching unit is connected via at least two switch communication links to other switching units and wherein each data receiver is connected to at least two different switching units. Subdividing the switching assembly in this way provides improved performance if any one of the switching units should fail. In a preferred embodiment, exactly one switching unit is attributed to each data source and, preferably, one input of each switching unit is connected an output of its data source.

In a further preferred embodiment, a synchronous transmission scheme is used where repetitive time windows are attributed to each data source. In each time window, the switching assembly connects all receiver communication links to the data source attributed to the time window. This again leads to an increase of required bandwidth, but it helps to keep the system simple. In addition to this, it prevents a faulty data source from jamming a receiver communication link continuously because each data source only has access to the communication link during its data window. If the switching assembly is divided into switching units connected as mentioned above, the same scheme can be used for preventing a jamming of the switch communication links.

In another preferred embodiment, a unique key is attributed to each data source. The data items sent by each source are digitally signed by the corresponding key, and the signature is checked upon receipt of the data item in a data receiver. Using such a signature scheme provides a further possibility for detecting corrupted messages.

The present invention is particularly suited for controlling the direction and/or velocity of vehicles. In a preferred application, it is used for controlling an aircraft.

Other preferred embodiments are described in the dependent claims.

The invention will be better understood and objects other than those set forth above will become apparent when consideration is given to the following detailed description thereof. Such description makes reference to the annexed drawings, wherein:

Fig. 1 is a block diagram of a fault tolerant computer system according to the present invention,

Fig. 2 shows a switching unit of one data source,

Fig. 3 is a timing schedule for data communication,

Figs. 4A, 4B, 4C are tables of redundant data items received by a data receiver,

Fig. 5 shows an aircraft controlled by a computer system according to the present invention,

Fig. 6 is a simplified illustration of Fig. 1 and

Fig. 7 is an alternative embodiment of the invention.

The system of Fig. 1 is operated by three redundant computers P0, P1 and P2, which process the signals from three sensors S0, S1 and S2 and two input devices V0, V1 and control three actuators A0, A1 and A1. One switching unit SU0, SU1, SU2 is attributed to each computer.

In a specific embodiment, the system shown here may control a vehicle, where the sensors S0, S1, and S2 e.g. measure the vehicle's position, attitude and/or velocity, the input devices V0 and V1 are controls operated by the user, and A0, A1 and A2 are actuators controlling the vehicle's drive and steering mechanism.

For reliability reasons, there are at least two redundant sensors for measuring each parameter used by the computers P0, P1, P2, and the input devices LV0, LV1 are provided in duplicate.

The computers P0, P1 and P2 generate commands for the actuators A0, A1, A2 as a function of the input provided by the sensors S0, S1 and S2 and the input devices LV0, LV1 as well as of state variables stored in the computers. The computers P0, P1 and P2 work independently. They are redundant, i.e. the commands generated by the computers are, in the absence of a system fault, ideally identical and therefore redundant. The commands are sent as data items to the actuators A0, A1, A2. The latter combine the received redundant data items in order to determine an error corrected data item. This is described in more detail below.

It must be noted that for most applications the corresponding number of components will be larger than shown in Fig. 1. However, the architecture of the present system can be scaled easily to meet the requirements of systems of any complexity.

In the following, this architecture is described in more detail. In this description, each computer P0, P1, P2 (or, more accurately, its output connected to the switching unit) is regarded as a "data source" sending data items to be received by the actuators A0, A1, A2. Each actuator A0, A1, A2 is regarded as a "data receiver" receiving the data items.

A plurality of communication links is provided for connecting the individual parts of the system. Input communication links LS0, LS1, LS2, LV0, and VL1 connect each sensor S0, S1, S2 and each input device V0, V1 to each computer P0, P1, P2. Switch communication links LPiPj interconnect the individual switching units SU0, SU1, SU2 (where i and j are integers between 0 and the number of switching units minus one). Receiver communication links LPiAk connect each switching unit SUi to the data receivers Ak (where k is an integer between 0 and the number of actuators minus 1). Each data receiver Ak is connected to at least two receiver communication links LPiAk leading to different switching units SUi.

Each switch communication link LPiPj is a point to point connection and connects one output of a switching unit SUi to one input of another switching unit SUj. Similarly, each receiver communication link LPiAk is a point to point connection connecting one output of a switching unit Si to one actuator Ak.

Preferably, the receiver communication links LPiAk are optical cables for reliable data transmission and safe galvanic protection of the remaining system because in many applications the actuators will operate high power equipment. The other communication links may be optical fibers, electric wires or radio links or others.

The architecture of the switching units SU0, SU1, SU2 is illustrated in Fig. 2. In the shown embodiment each switching unit SUi has three inputs I0 - I2 and five outputs O0 - O4. One switch (demultiplexer) S0 - S4 is provided for each output so that each output Oi can be selectively connected to any one of the inputs Ij.

Inputs I0 and I2 are each connected to a switch communication link LPjLPi, LPj'LPi receiving data items from two other switching units SUj and SUj'. Input I1 is connected to the data source attributed to the switching unit.

Outputs O0 and O4 are each connected to a switch communication link LPiLPj, LPiLPj' for sending data items to two other switching units SUj and SUj'. Outputs O1 and O2 are connected to receiver communication links LPiAk and LPiAk' for sending data items to receivers Ak and Ak'. Output O3 is connected to a data input of the computer attributed to the switching unit.

A switch control table 10 is provided for setting the switches Si in accordance with signals from a clock unit 11.

Each switching unit SU0, SU1, SU2 is provided with its own clock unit 11 and its own table 10 in order to be able to set the switches autonomously. The clock

units 11 are kept synchronized. Various fault tolerant methods for keeping clocks synchronized are known to the person skilled in the art, some of which are described by Fred. B. Schneider in "Understanding Protocols for Byzantine Clock Synchronization", August 1987, Dept. of Computer Science, Cornell University. Preferably, the clock units 11 are synchronized by time stamps added by the data sources to each or at least part of the data items, wherein each switching unit extracts the time stamp of passing data items from different data sources and determines a global time therefrom, e.g. by finding a median of the time stamps received at one time and by calculating a deviation in respect to its own clock.

For regulating communication in the system, a time window is attributed to each data source, wherein the windows are preferably of equal length and are repeated at regular cycles as shown in Fig. 3. Data windows of unequal length may also be used, in particular if one of the data sources has a larger amount of data to transmit. In a given time window, the switching units SUi set the switches in such a way that all switch communication links LPiPj as well as all receiver communication links LPiAk are connected to the data source the window is attributed to.

As can be seen from Fig. 3 and as will be explained further below, additional time windows may be provided for transmissions from the actuators Ai.

The data sources are also being synchronized, e.g. through the clock units of their attributed data switches, and only send data items within their data windows, wherein a leading and trailing end of each data window remains unused in order to account for synchronization mismatch and signal delays.

The lengths of the windows in Fig. 3 primarily depends on the amount of data to be transported and the maximum allowable time delay for transmitting a mes-

sage. For most vehicle control systems, a window length in the order of 10 ms is found to be appropriate.

Using a fixed timing scheme for globally attributing the communication links to a single data source at a time leads to an increase in bandwidth requirements. However, in many applications, presently available communication links provide ample bandwidth for supporting this type of protocol.

As it becomes clear from the above, each data source Pi sends all its data items to all data receivers Ak simultaneously, and each data receiver receives every data item through at least two different receiver communication links LPjAk simultaneously. Hence, in normal operation, the data receiver receives each data item from each data source at least twice, and because all data sources are generating redundant data items, the data receiver receives a group of six redundant versions of each data item through different paths of the network.

This is illustrated for data receiver A0 in Figs. 4A, 4B and 4C. The data receiver tries to receive all six data items of the group and can verify their physical integrity, e.g. by verifying a check sum or a digital signature as described below. In the absence of any error in transmission, each data item is flagged as "ok" as shown in Fig. 4A. In case of a failure of communication link LP1P0, only five data items are valid, Fig. 4B. Even if, in addition to this, communication link LP2A0 fails, two data items are still valid, Fig. 4C.

From the valid received redundant data items, the data receiver generates an error corrected data item using known permutation-invariant techniques (median, majority, ...). For example, if the data items specify a numerical parameter, the median value of the parameter given by the valid data items is determined.

As mentioned above, the data items can comprise a digital signature. In order to generate a digital signature (and, optionally, an encryption), a unique key

is attributed to each data source P0, P1, P2. Using this unique key, each data source creates a digital signature as known to a person skilled in the art, i.e. a signature value that depends on the message to be transmitted in the data item as well as on the key, wherein the algorithm used for generating the signature is such that it is possible to verify with sufficient reliability if a given signature value was generated using a given key or not. For improved security, signature schemes based on asymmetric keys can be used. It must be noted, however, that the signature schemes that can be used in the context of the present invention may be simpler and less tamper-proof than those generally used in data communication because they primarily have to protect against system failure but not against intentional tampering.

When a data receiver receives a message from a given data source, it checks the validity of the data item by checking if the signature matches the key of the data source. If not, the data item is flagged to be invalid.

An application of the present system is schematically illustrated in Fig. 5. The figure shows a VTOL aircraft 20 as it is e.g. disclosed in WO 01/30652 with a plurality tiltable drive units 21, each of which comprises an electrically driven fan. The drive units 21 provide attitude control, lift and forward thrust for the aircraft. Each drive unit 21 comprises a drive control unit for controlling its tilt angle and thrust. Each control unit receives its settings from one of the data receivers Aj, Ai described above. In addition to this, attitude sensors Sm, Sn and other types of sensors as well as the input devices V0 and V1 are arranged in the aircraft for providing the computers Pi with input data.

In order to discuss some of the many modifications of the present invention, we now refer to Fig. 6, which shows the embodiment of Fig. 1 in schematic manner.

As can be seen from Fig. 6, one of the advantages of the described embodiment of the present invention lies in the fact that each data receiver Ak receives data from all data sources Pi over redundant paths even though the number of receiver communication links LPiAk for a given receiver Ak is smaller (namely 2) than the number of data sources (namely 3). This is due to the fact that the switching units SUi allow each data source Pi to access both receiver communication links of a given data receiver.

The minimum number of receiver communication links to each data receiver is 2 if alternative paths are to be provided for each data item. In order to increase reliability, more than two receiver communication links for each data receiver could be provided.

In the embodiment of Fig. 6, each switching unit SUi is connected for sending and receiving data with two other switching units, thereby providing alternative paths between switching units. For increased reliability, this number can be larger than two, but there may also be only one single switching communication link per switching unit.

In the embodiment of Fig. 6, one switching unit SUi is attributed to each computer Pi. Preferably, each computer and each switching unit are located physically close to each other such that they may share some mechanical or electrical components. However, it is preferred that the switching unit is able to operate independently of its attributed computer, i.e. when the computer fails in its data processing, the switching unit should still continue to operate.

A close physical placement of the computer and its associated switching unit is preferred but not required. The switching unit can be placed at an arbitrary position. However, if the distance between a computer and its switching unit becomes large, the risk of failure of the communication links between them in-

creases. In that case it can be advisable to provide an additional redundant communication links between the computer and the switching units.

Fig. 7 illustrates an embodiment with four computers Pi and only two switching units SUj. Here, each switching unit has four inputs and six outputs, and the individual switches S have four possible positions. Again, the switches are positioned according to the data source the current window is attributed to such that the signals of this data source are sent to all receiver communication links and to the switching communication links.

It must be noted that in the above description and the enclosed figures, only the most important ones of the communication links between the components are described and shown. In addition to this, the network may comprise further communication links, e.g. from the actuators back to the computers or to a separate monitoring unit. Similarly, the switching unit may comprise, in addition to switches connected to the receiver communication links and the switching communication links, additional switches for feeding data to other types of receivers, such as the switches S3 for feeding data items to the inputs of the computers.

For example, if it is desired that the actuators A0, A1, A2 are able to send feedback to the computers P0, P1, P2, the communication links between the switching units SUi and the actuators may be bidirectional. For example, a feedback link LAiPk may lead from each actuator to the two switching units it is connected to, such as it is shown in dashed lines for actuator A1 in Fig. 1. The number of inputs to the switches S0 - S4 in each switching unit would correspondingly be increased by two, such that they are able connect the feedback links LAiPk to the outputs of the switching units during time windows attributed to the actuators (see Fig. 3). In other words, actuators A0, A1, A2 can also act as data

sources. However, in contrast to computers P0, P1 and P2, they are generally not redundant data sources, but they can transmit the data over redundant paths to the selected receivers.

The term "system" as used here is understood to designate an apparatus comprising data sources and data receivers as well as the network connecting them, but it is also used to designate a method for operating such an apparatus.

Claims

1. An error tolerant computer controlled system comprising

a plurality of redundant data sources (P0, P1, P2) generating at least partially redundant data items,

a plurality of data receivers (A0, A1, A2) for receiving the redundant data items and combining them to an error tolerant data item,

a switching assembly (SU0, SU1, SU2) with a plurality of inputs and outputs, wherein each input is connected to one data source (P0, P1, P2) or to one output and wherein each output is connected to one input or to one data receiver (A0, A1, A2), and wherein each data receiver (A0, A1, A2) is connected via separate receiver communication links (LPiAk) to at least two outputs,

wherein said switching assembly (SU0, SU1, SU2) is adapted to connect any of said data sources (P0, P1, P2) to each of said data receivers (A0, A1, A2) over at least two different receiver communication links (LPiAk), and wherein said computer controlled system is adapted to send every data item from any given data source (P0, P1, P2) to any given data receiver (A0, A1, A2) through every one of the at least two different receiver communication links (LPiAk) such that the given data receiver receives the same data item through at least two receiver communication links (LPiAk).

2. The system of claim 1 wherein each receiver communication link (LPiAk) connects exactly one output to exactly one receiver.

3. The system of any of the preceding claims wherein the number of receiver communication links (LPiAk) for each data receiver (A0, A1, A2) is smaller than the number of data sources (P0, P1, P2), and in par-

ticular wherein the number of receiver communication links (LPiAk) for each data receiver (A0, A1, A2) is 2.

4. The system of any of the preceding claims wherein the switching assembly (SU0, SU1, SU2) is divided into a plurality of switching units, wherein each input of each switching unit is either connected to one data source (P0, P1, P2) or via a switch communication link (LPiPj) to one output of another switching unit, wherein each switching unit is connected via at least two switch communication links (LPiPj) to other switching units, wherein each switch communication link (LPiPj) connects one output to one input, and wherein each data receiver (A0, A1, A2) is connected via the receiver communication links (LPiAk) to at least two different switching units (SU0, SU1, SU2), and in particular wherein, for each switching unit, each output can be connected to each input.

5. The system of claim 4 wherein exactly two switch communication links (LPiLPj) are attached to the inputs of each switching unit (SUi) and/or wherein exactly two switch communication links (LPiLPj) are attached to the outputs of each switching unit (SUi).

6. The system of any of the claims 4 to 5 wherein the number of switching units (SUi) corresponds to the number of data sources (P0, P1, P2) and wherein each switching unit is attributed to one data source (P0, P1, P2), and wherein one input of each switching unit is connected to its attributed data source (P0, P1, P2), and in particular wherein one output of the switching unit is connected to its attributed data source (P0, P1, P2).

7. The system of any of the preceding claims wherein repetitive time windows are attributed to each data source (P0, P1, P2) and wherein, in each time window, the switching assembly (SU0, SU1, SU2) connects all receiver communication links (LPiAk) to the data source (P0, P1, P2) attributed to the time window while discon-

necting the remaining data sources (P0, P1, P2) from the receiver communication links (LPiAk).

8. The system of claim 7 and of any of the claims 4 to 6 wherein, in each time window, the switching assembly (SU0, SU1, SU2) connects all switch communication links (LPiPj) to the data source (P0, P1, P2) attributed to the time window while disconnecting the remaining data sources (P0, P1, P2) from the switch communication links (LPiPj).

9. The system of any of the claims 7 to 8 wherein at least part of the data items carries a time stamp and wherein each switching unit (SUi) comprises a clock (11) synchronized by the time stamps.

10. The system of claim 9 wherein each switching unit (SUi) is adapted to combine a plurality of received data items carrying time stamps in order to determine a time base, in particular by determining a median of the time stamps of data items from different data sources (P0, P1, P2).

11. The system of any of the preceding claims wherein a unique key is attributed to each data source (P0, P1, P2) and each data source (P0, P1, P2) is adapted to generate a digital signature for each data item it sends using its unique key, and wherein the data receivers (A0, A1, A2) are adapted to check a validity of the signature upon receipt of a data item.

12. The system of any of the preceding claims wherein said data receiver (A0, A1, A2) is adapted to check a validity of each of the received data items and to use only those data items of a group of redundant data items that are valid, and in particular wherein it determines a median or majority value of the valid data items of the group of redundant data items.

13. The system of any of the preceding claims wherein said data receivers (A0, A1, A2) comprise actuators.

14. The system of any of the preceding claims further comprising feedback links (LAiPk) for transmitting data from said data receivers to said switching assembly.

15. A vehicle comprising the system of any of the preceding claims, wherein said data receivers (A0, A1, A2) control a drive and steering mechanism of the vehicle.

16. An aircraft comprising the system of any of the claims 1 to 14.

17. The aircraft of claim 16 comprising at least one pivotal drive unit (21) for attitude control and for generating lift and forward thrust, and a drive control unit for controlling a tilt angle and a thrust of said drive unit, wherein said control unit is controlled by one of said data receivers (A0, A1, A2), and in particular wherein said drive unit is driven by an electrical motor.

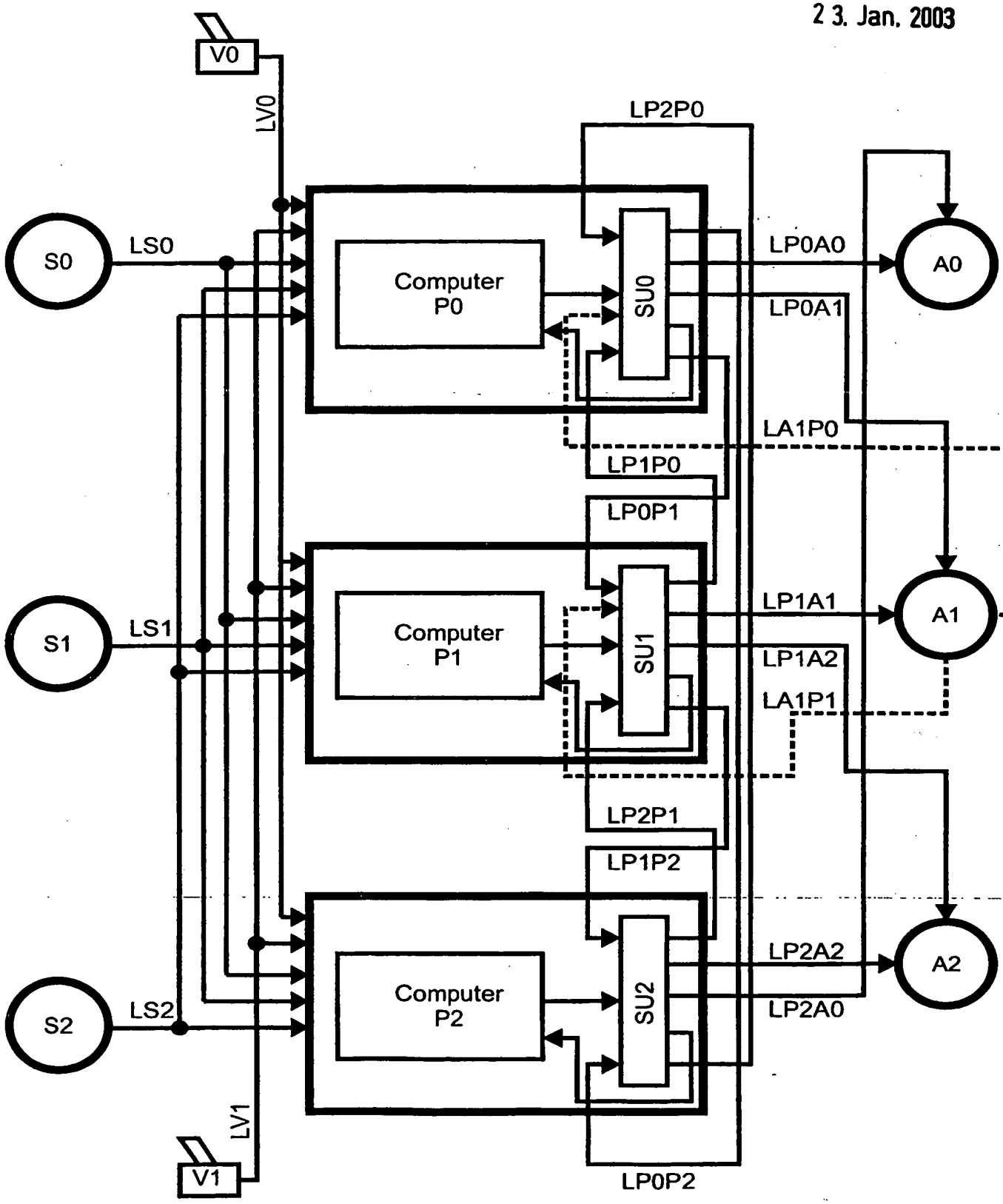

Fig. 1

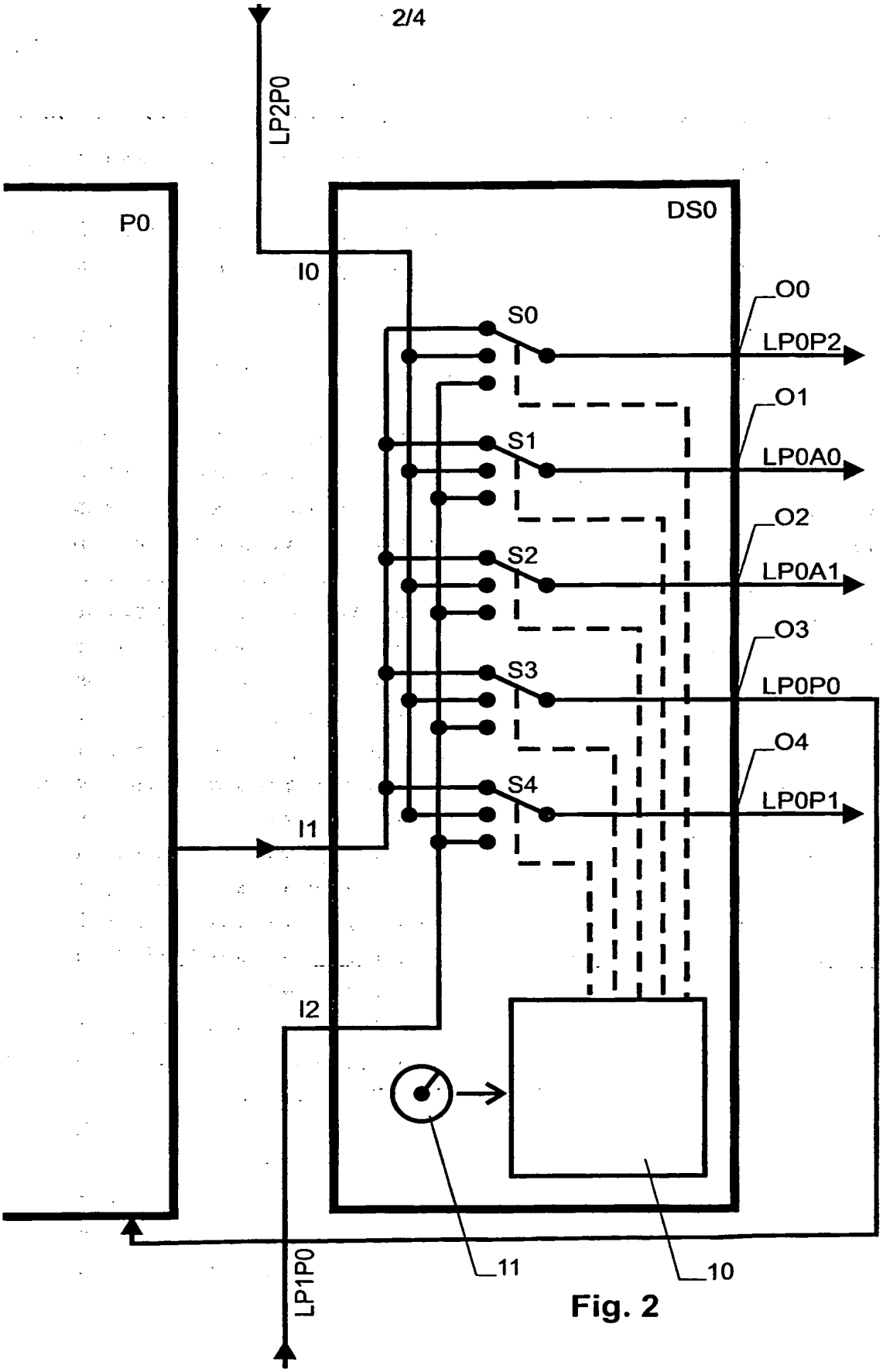
LP2P0
P0
DS0
I0
S0
O0
LP0P2
S1
O1
LP0A0
S2
O2
LP0A1
S3
O3
LP0P0
S4
O4
LP0P1
I1
I2
11
10
LP1P0

Fig. 2

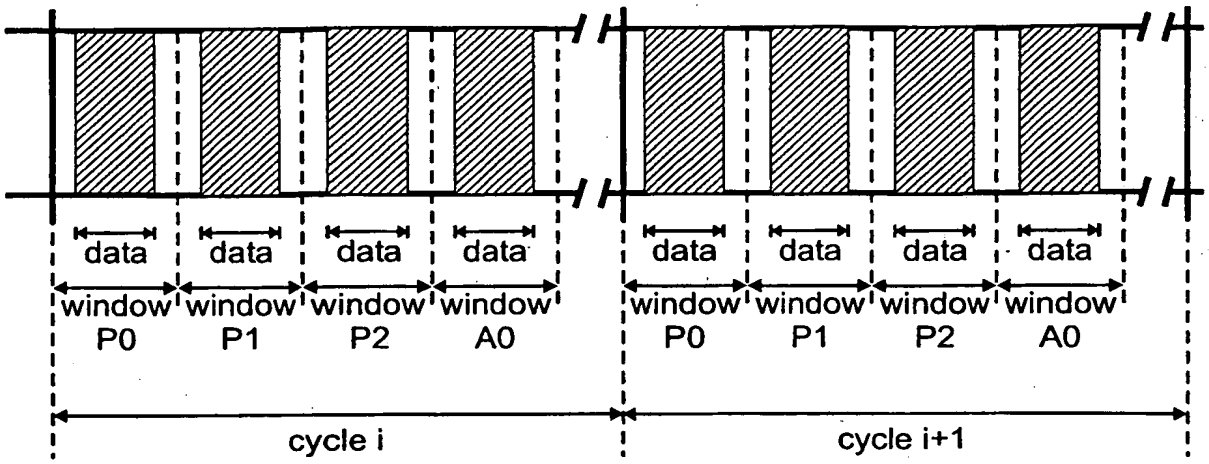

Fig. 3

	through LP0A0	through LP2A0
from P0	ok	ok
from P1	ok	ok
from P2	ok	ok

Fig. 4A

	through LP0A0	through LP2A0
from P0	ok	ok
from P1		ok
from P2	ok	ok

Fig. 4B

	through LP0A0	through LP2A0
from P0	ok	
from P1		
from P2	ok	

Fig. 4C

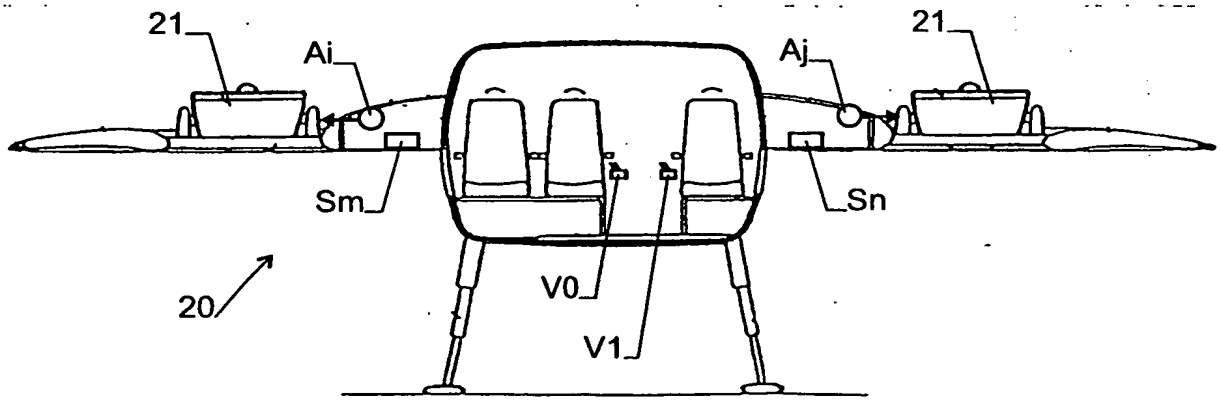

Fig. 5

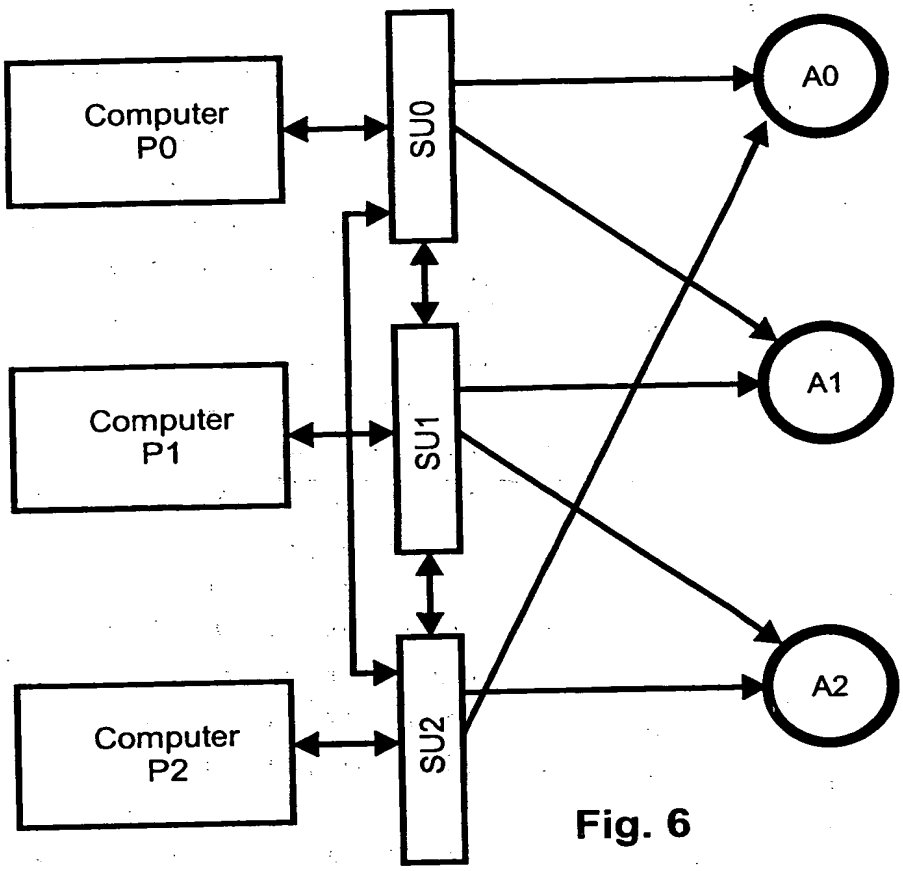
Computer
P0
Computer
P1
Computer
P2
SU0
SU1
SU2
A0
A1
A2

Fig. 6

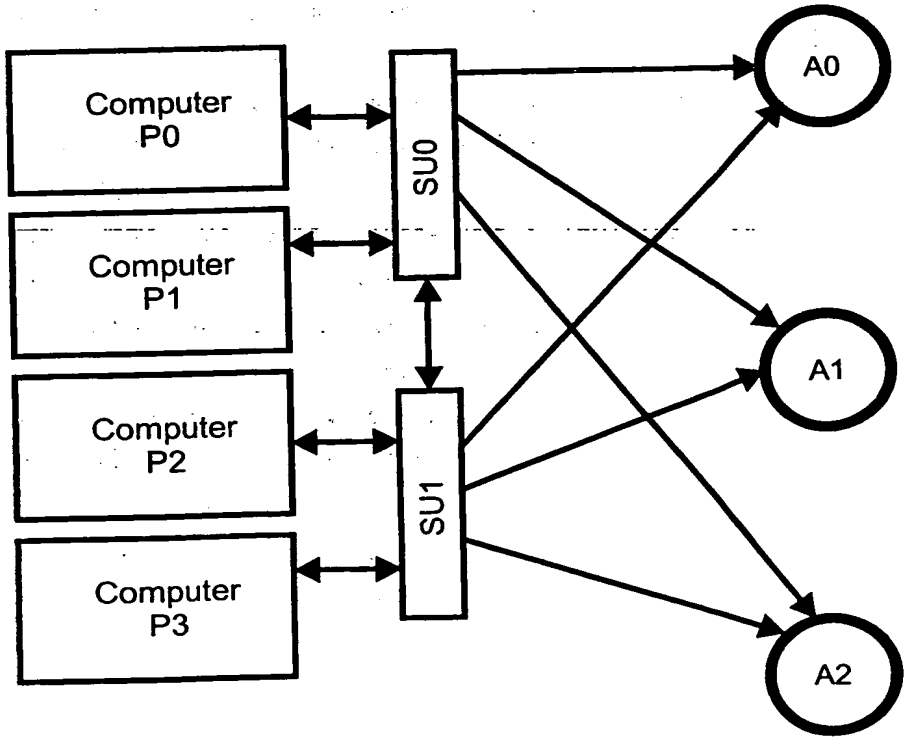
Computer
P0
Computer
P1
Computer
P2
Computer
P3
SU0
SU1
A0
A1
A2

Fig. 7

Abstract

An error tolerant computer controlled system comprises several computers (P0, P1, P2) working redundantly and controlling actuators (A0, A1, A2) based on signals from sensors (S0, S1, S2) and input devices (V0, V1). Each data item emitted by each computer is simultaneously sent through differing communication paths to each actuator, such that in normal operation each actuator receives each data item through several paths. This system continues to function properly even in case of a failure without requiring any re-routing of the data items, which makes it easier to design, analyze and test and thereby increases its reliability.

(Fig. 1)